FINAL REPORT

Capstone Project - The Battle of Neighbourhoods

Introduction:

New York is the largest city in the United states with a lot of immigration. The population of New York city was close to 8.5 million in 2014. The large portion of these immigrants includes Chinese, Indians and French. This is a hypothetic project to identify the potential clusters of area having Indian restaurants to determine which areas would be better to open new Indian restaurants.

By using data science methods and tools along with the machine learning algorithms such as clustering, the project aims to provide solution to answer few business considerations.

Problem:

To find the answers to the following questions:

Q1) List and visualize all major parts of New York City that has great Indian restaurants.

Q2) what is best location in New York City for Indian Cuisine?

Q3) which areas have potential Indian Restaurant Market?

Q4) which all areas lack Indian Restaurants?

Q5) which area is preferred for the Indian Cuisine?

Q6) Which area should be selected for the opening of Indian restaurant.

Data Section:

New York City's demographics show that it is a large and ethnically diverse metropolis. With its diverse culture, comes diverse food items. There are many restaurants in New York City, each belonging to different categories like Chinese, Indian, and French etc.

For this project we need the following data:

- New York City data that contains list Boroughs, Neighbourhoods along with their latitude and longitude.
  - Data source : https://cocl.us/new_york_dataset
  - Description: This data set contains the required information. And we will use this data set to explore various neighbourhoods of New York City

- Indian restaurants in each neighbourhood of New York City.
  - Data source : Foursquare API
  - Description: By using this API we will get all the venues in each neighbourhood. We can filter these venues to get only Indian restaurants.

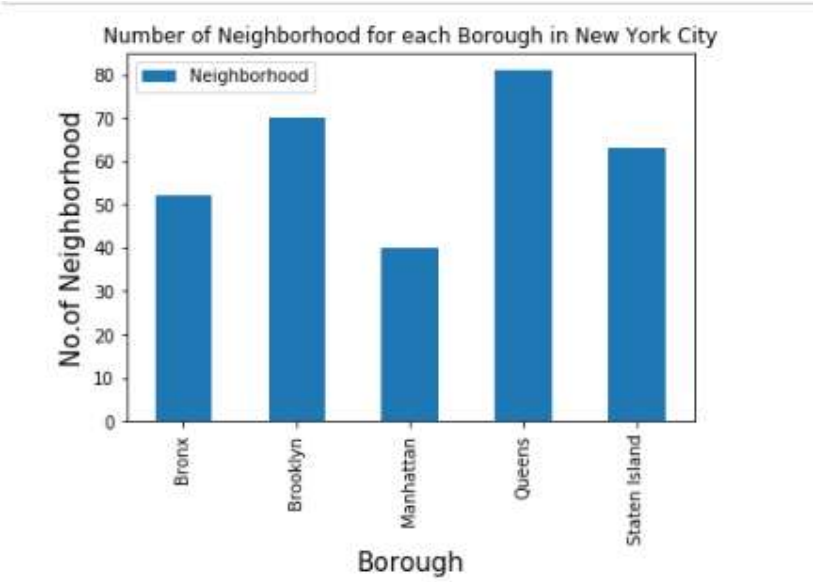
Methodology:

1. We begin by collecting the New York city data from the following link "https://cocl.us/new_york_dataset"
2. We will find all venues for each neighbourhood using Foursquare API

```
: new_york_data=get_new_york_data()
new_york_data.head()
```

[6]:

	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585



3. We will then filter out all venues with Indian restaurant for further analysis.

```
( 1 / 306 ) Indian Resturants in Wakefield, Bronx:0
( 2 / 306 ) Indian Resturants in Co-op City, Bronx:0
( 3 / 306 ) Indian Resturants in Eastchester, Bronx:0
( 4 / 306 ) Indian Resturants in Fieldston, Bronx:0
( 5 / 306 ) Indian Resturants in Riverdale, Bronx:0
( 6 / 306 ) Indian Resturants in Kingsbridge, Bronx:0
( 7 / 306 ) Indian Resturants in Marble Hill, Manhattan:0
( 8 / 306 ) Indian Resturants in Woodlawn, Bronx:1
( 9 / 306 ) Indian Resturants in Norwood, Bronx:0
( 10 / 306 ) Indian Resturants in Williamsbridge, Bronx:0
( 11 / 306 ) Indian Resturants in Baychester, Bronx:0
( 12 / 306 ) Indian Resturants in Pelham Parkway, Bronx:0
( 13 / 306 ) Indian Resturants in City Island, Bronx:0
```

4. Next using Foursquare API, we will find the Ratings, Tips, and Number of Likes for all the Indian Restaurants.

	Borough	Neighborhood	ID	Name	Likes	Rating	Tips
0	Bronx	Woodlawn	4c0448d9310fc9b6bf1dc761	Curry Spot	5	8.1	10
1	Bronx	Parkchester	4c194631838020a13e78e561	Melanies Roti Bar And Grill	3	5.9	2
2	Bronx	Spuyten Duyvil	4c04544df423a593ac83d116	Cumin Indian Cuisine	13	5.8	9
3	Bronx	Concourse	551b7f75498e86c00a0ed2e1	Hungry Bird	8	7.0	3
4	Bronx	Unionport	4c194631838020a13e78e561	Melanies Roti Bar And Grill	3	5.9	2

5. We will then sort Neighbourhoods and Borough the data keeping Ratings as the constraint.

	Neighborhood	Average Rating
5	Blissville	9.00
61	Sunnyside	9.00
12	Civic Center	8.90
70	West Village	8.90
29	Greenwich Village	8.90
65	Tribeca	8.90
23	Fort Greene	8.80
49	Prospect Heights	8.70
0	Astoria	8.65
63	Sutton Place	8.50

	Borough	Average Rating
2	Manhattan	8.239394
1	Brooklyn	7.680769
3	Queens	6.609333
0	Bronx	5.671429
4	Staten Island	4.514286

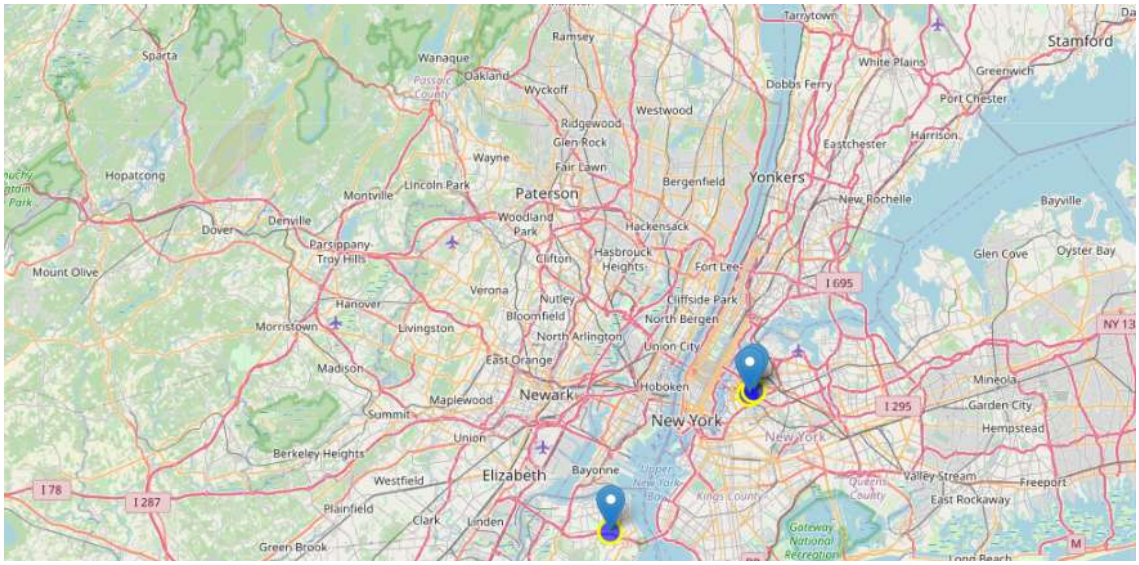
6. Next we will consider all the neighbourhoods with average rating greater or equal 9.0 to visualize on map.

	Borough	Neighborhood	Latitude	Longitude	Average Rating
0	Queens	Blissville	40.737251	-73.932442	9.0
1	Queens	Sunnyside	40.740176	-73.926916	9.0
2	Staten Island	Sunnyside	40.612760	-74.097126	9.0

7. We will join this dataset to original New York data to get longitude and latitude.

8. Finally, we will visualize the Neighbourhoods based on average Rating using python's Folium library.

   Neighbourhoods based on average rating:



Result:

So now we can answer the questions asked above in the Questions section:

Answers:

1. The following location in New York City has great Indian restaurants.

	Borough	Neighborhood	Latitude	Longitude	Average Rating
0	Queens	Blissville	40.737251	-73.932442	9.0
1	Queens	Sunnyside	40.740176	-73.926916	9.0
2	Staten Island	Sunnyside	40.612760	-74.097126	9.0

2. Bliss ville and Sunnyside are some of the best neighbourhoods for Indian cuisine.
3. Manhattan have potential Indian restaurant market because of having highest average ratings.
4. Staten Island ranks last in average rating of Indian restaurants.
5. Manhattan is a preferred area for Indian restaurants.

Conclusion:

If you open restaurant in Manhattan, you will be having a tough competition as most of the restaurants are having high rating and you will have to compete with those. If you open a restaurant in Staten Island, because of the average ratings, if one keeps high quality and provides better service with variety of dishes, the chances are high to have the success as the ratings are very poor.

Apart from the above analysis, one should look for the population of Indian origin in this area to determine the population of potential customers in these areas.